บทที่ 1

บทนำ

ความเป็นมา

การสังเคราะห์เสียงพูดมีแนวโน้มกระทำโดยการคำนวณด้วยเครื่องคอมพิวเตอร์ตั้งแต่ปี ค.ศ. 1970 เป็นต้นมา การคำนวณเป็นการประมวลผลสัญญาณในลักษณะเชิงเลข (Digital Signal Processing) ซึ่งต้องอาศัยอุปกรณ์ที่มีความสามารถในการคำนวณสูง เกี่ยวกับทฤษฎีที่ใช้ในการวิเคราะห์และสังเคราะห์เสียงพูดระหว่างปี ค.ศ. 1969-1978 วิธีการเข้ารหัสแบบลิเนียร์พรีดิกทีฟ (Linear Predictive Coding) หรือแอลพีซีได้รับความนิยมมากที่สุด ในช่วงเวลาดังกล่าวยังไม่มีอุปกรณ์การสังเคราะห์เสียงพูดใดที่สามารถใช้งานในเชิงพาณิชย์ได้ จนถึงปี ค.ศ. 1980 บริษัท Texas Instruments ได้ผลิตวงจรรวมสำหรับการสังเคราะห์เสียงพูดด้วยวิธีแอลพีซีซึ่งสามารถนำไปใช้ในอุตสาหกรรมได้ทันที แต่ต้นทุนยังสูงและมีการปกปิดทางเทคโนโลยีทำให้ผู้ใช้ไม่สามารถนำวงจรรวมมาใช้งานได้เอง ในระยะเวลานั้นเองก็มีอุปกรณ์สังเคราะห์เสียงพูดผลิตออกสู่ตลาดมากมายจนถึงระดับ Text to Speech แต่ยังคงเป็นภาษาอังกฤษทั้งสิ้น เมื่อปลายปี ค.ศ. 1982 บริษัท Texas Instruments ได้ผลิตวงจรรวมประมวลผลสัญญาณที่เรียกว่า Signal Processor [1] วงจรรวมดังกล่าวมีลักษณะเป็นไมโครโปรเซสเซอร์ที่มีความสามารถในการคำนวณสูง ใช้งานสะดวก เนื่องจากอุปกรณ์ดังกล่าวสามารถกำหนดการทำงานโดยผู้ใช้ได้ทั้งหมด หรือการทำงานขึ้นอยู่กับซอฟท์แวร์ซึ่งผู้ใช้เป็นผู้เขียนขึ้น อย่างไรก็ตามการจะผลิตอุปกรณ์สำหรับสังเคราะห์เสียงพูด จำเป็นต้องมีระบบที่ใช้สำหรับการทดลองในการศึกษาเกี่ยวกับเสียงพูด ระบบซึ่งสามารถเชื่อมโยงระหว่างทฤษฎีกับอุปกรณ์ที่ทำงานได้จริง ดังนั้นโครงงานจึงต้องเน้นหนักที่ระบบวิเคราะห์-สังเคราะห์เพื่อที่จะได้เป็นรากฐานการพัฒนางานวิจัยในเรื่องของเสียงพูดต่อๆ ไป

วัตถุประสงค์

1. เพื่อการศึกษากรรมวิธีในการสังเคราะห์เสียงพูดและการทดลองการสังเคราะห์เสียง

2. ออกแบบสร้างระบบที่สามารถวิเคราะห์และสังเคราะห์เสียงพูด ซึ่งจะเป็นแนวทางสำหรับการศึกษาเกี่ยวกับเสียงพูดในเชิงวิศวกรรม

ขอบเขตของงานวิจัย

1. ทำการศึกษาและทดลองเพื่อหากรรมวิธีที่เหมาะสมและมีประสิทธิภาพสำหรับการสังเคราะห์เสียงพูดเป็นคำๆ โดยอาศัยทฤษฎีการเข้ารหัสแบบลิเนียร์พรีดิกทีฟ

2. ทดลองวิเคราะห์และสังเคราะห์เสียงพูดภาษาไทยเป็นคำๆ โดยการคำนวณในเครื่องคอมพิวเตอร์ ในการนี้ได้กำหนดตัวอย่างคำขึ้นมาเป็นตัวเลข "หนึ่ง" ถึง "สิบ"

3. สร้างอุปกรณ์เฉพาะสำหรับเปล่งเสียงคำพูดทีละคำทันทีทันใด โดยการคำนวณตามกรรมวิธีเข้ารหัสแบบลิเนียร์พรีดิกทีฟ

ขั้นตอนและวิธีดำเนินการวิจัย

1. คัดเลือกกรรมวิธีที่เหมาะสมในการสังเคราะห์เสียงพูดเป็นคำๆ โดยคำนึงถึงคุณภาพและประสิทธิภาพในการสังเคราะห์เสียง ขั้นตอนนี้อาศัยผลการทดลองจากการคำนวณด้วยเครื่องคอมพิวเตอร์

2. ออกแบบและสร้างเครื่องต้นแบบรวมทั้งอุปกรณ์สนับสนุนต่าง ๆ ซึ่งได้แก่ภาคแปลงสัญญาณระหว่างสัญญาณอนาลอกกับสัญญาณเชิงเลข ภาคประมวลผลสัญญาณและภาคอินเตอร์เฟซกับระบบไมโครคอมพิวเตอร์

3. ปรับปรุงกรรมวิธีในข้อ 1. ในส่วนของการสังเคราะห์เสียงเพื่อประยุกต์ใช้กับภาคประมวลผลสัญญาณซึ่งอาศัยการทำงานของไมโครโปรเซสเซอร์ TMS 32010

4. ทำการวิเคราะห์หาข้อมูลที่ใช้ในการสังเคราะห์เสียงด้วยวิธีเข้ารหัสแบบลิเนียร์พรีดิกทีฟของตัวอย่างคำพูด

5. ทดสอบการทำงานและปรับปรุงคุณภาพของเสียง รวบรวมข้อมูลและประเมินผล

<u>ประโยชน์ที่จะได้รับจากการวิจัยนี้</u>

1. สามารถนำไปประกอบกับอุปกรณ์ต่าง ๆ เพื่อให้มีการตอบรับเป็นเสียงพูดภาษาไทย

2. สามารถประยุกต์เพื่อใช้งานเกี่ยวกับการส่งสัญญาณเสียงพูดในระบบสื่อสารในแง่ต่าง ๆ เช่น ลดทอนจำนวนข้อมูลเพื่อเพิ่มความหนาแน่นของข้อมูล ปรับปรุงคุณภาพเสียงที่สูญเสียระหว่างการสื่อสาร และการทำ Voice Scramble เพื่อป้องกันการรั่วไหลของข่าวสาร เป็นต้น

3. กรรมวิธีวิเคราะห์-สังเคราะห์เสียงและระบบที่สร้างขึ้น สามารถนำไปใช้ศึกษาวิจัยในเรื่องต่าง ๆเกี่ยวกับเสียงพูดทางด้านวิศวกรรมเช่น

- การรู้จำเสียงพูด (Speech Recognition)

- การเพิ่มคุณภาพของสัญญาณเสียงพูด (Speech Signal Quality Enhancement)

- การพิสูจน์เสียงพูด (Speech Verification)

นอกจากนั้นยังสามารถนำไปใช้ประโยชน์เพื่อการศึกษาภาษาศาสตร์ในประเทศ

4. ในส่วนของระบบประมวลผลสัญญาณที่สร้างขึ้น สามารถนำไปใช้พัฒนาการใช้งานของวงจรรวมประมวลผลสัญญาณ TMS 32010